Remarks/Arguments:

Claims 1-9, 12, 16-20, 22-27, 30, 35, 37, 38, 40, and 47-50 are currently pending. Claims 1, 7, 8, 12, 16, 19, 23, 37, and 48-50 have been amended for clarification and are supported by the original claims and Figures 3A and 3B and paragraphs 0051-0054. Claims 51-57 have been added to enhance the scope of Applicant's patent coverage and are supported by paragraph 0060 and Figure 3B of the application as filed. It is respectfully submitted that no new matter has been added.

Response to two arguments on pages 2-3 of the Final Office Action dated August 21, 2008.

As to the first argument, the claim language has been changed so the argument is moot. The claims have been amended to relate to distributed feature extraction in which an apparatus that is configured to perform lower level but not higher level feature extraction and a remote service that is configured to perform higher level feature extraction from extracted lower level features transmitted by the apparatus, subject matter not believed to be found or obvious from the cited references.

As to the second argument, Applicant's claimed invention relates to 1) an apparatus configured to extract lower level features that may be later used remotely for extracting higher level features off-apparatus from the lower level features to identify a media from a media sample and 2) a remote service configured to use received lower level features and configured, if needed, to extract higher level features from the received lower level features to uniquely identify a media corresponding to the media sample. The "client sends a feature-extracted summary of the captured signal sample containing landmark and fingerprint pairs to the server end" in Wang (column 8, lines 13-15). Barton discloses samples being from an experiential environment 101 being sent over a network to a recognition engine 110 that derives sufficient characteristics to enable a predetermined event to be triggered 130 (paragraph 0048). In his abstract, Barton discloses the triggered events include the delivery of information and services to the user, the execution of tasks and instructions by the service on the user's behalf, communication events, surveillance events and other control-oriented events that are responsive to the user's wishes.

35 U.S.C. 112, second paragraph, rejection

The Patent Office rejected claim 49 under 35 U.S.C. 112, second paragraph, as being indefinite. Claim 49 has been amended to recite "apparatus" and not "mobile

station." Accordingly, Applicant respectfully requests that the Patent Office withdraw its rejection of claim 49 under 35 U.S.C. 112, second paragraph.

<div align="center">

35 U.S.C. 103(a) rejection

</div>

The Patent Office rejected claims 1-9, 12, 16-20, 22-27, 30, 35, 37, 38, 40, and 47-50 under 35 U.S.C. 103(a) as being unpatentable over Wang, U.S. Patent No. 6,990,453, in view of Barton, U.S. Published Patent Application No. 2002/0072982, and Vetro, U.S. Patent No. 6,490,320.

In Applicant's invention, as supported by paragraph 0050, there are an apparatus and a remote service that are configured to perform distributed feature extraction, wherein the apparatus is configured to perform lower level feature extraction and the remote service is configured to perform any needed higher level feature extraction from extracted lower level features transmitted by the apparatus to identify the media from which the lower level features have been extracted.

In Wang, landmarks and fingerprints are used to build a database 18. A media sample is captured 12 (Figure 1). Landmarks and fingerprints from the exogenous media sample are computed 14 and matched 16 through use of the database 18. Correspondences are generated 20 and a winning media sample file is located 22.

Wang discloses a sound source continually sampled into a buffer (column 21, lines 64-67). Sound parameters may be extracted from a sound buffer into fingerprints or other intermediate feature-extracted forms and stored in a second buffer (column 22, lines 19-21). New fingerprints may be added to the front of the second buffer while old fingerprints are discarded from the end of the buffer to form a rolling buffer (column 22, lines 22-24).

The method of Wang involves a search first performed on a first subset of sound files and only if the first search fails, then a search of second subset of sound files is performed (column 19, lines 23-34). Wang's method does not involve requesting the mobile station to provide a second set of features and does not appear amenable to modification to request a second set of features from the mobile station since the method of Wang involves a first search of highly used sound files only to be followed by a second search of less highly used sound files. Wang does not contemplate a request for a second set of features, as evidenced by Figure 1, in which Wang finds matching fingerprints 16 and then generates correspondences 20 with sample landmarks to find a winning sound file 22.

Furthermore, Wang does not teach an apparatus configured to extract lower level features that may be later used to extract higher level features off-apparatus to identify a media from a media sample and does not teach a remote service configured to identify from received lower level features and configured to extract higher level features from the received lower level features to uniquely identify a media corresponding to the media sample. Wang, in column 8, lines 13-24, discloses as follows:

> The client end sends a feature-extracted summary of the captured signal sample containing landmark and fingerprint pairs to the server end, which performs the recognition. Sending this feature-extracted summary to the server, instead of the raw captured signal, is advantageous because the amount of data is greatly reduced, often by a factor of 500 or more. Such information can be sent in real time over a low-bandwidth side channel along with or instead of, e.g., an audio stream transmitted to the server. This enables performing the invention over public communications networks, which offer relatively small-sized bandwidths to each user.

In Wang, the feature extraction is disclosed as occurring in the client device and the recognition occurs in the server. The computational nodes referenced in column 15, lines 12-14, in Wang correspond to the client side of Wang's system. On the server side, Wang, from column 15, line 59, through column 18, line 50, the extracted features are used to rank candidates; no features are extracted from these features received from the client device.

The Patent Office from page 3, line 14, through page 4, line 4, of the Office Action dated January 08, 2008, as follows:

> In an analogous art, Barton teaches a system for identifying audio samples that includes a recursive feature for automatically requesting more information in order to narrow the search results to find the corresponding file. (Page 5 [0048 and 0049] the "resolution of the derivation is coupled, in large measure, to the level of discrimination required in selecting an event to be triggered. As the number of potentially triggered events increases, the necessity to resolve ambiguity in the sample also increases," Page 6 [0059] "the song excerpt may be increased in length, or a different excerpt may be furnished, in an iterative manner" until a song is identified and Page 7 [0067-0068]) At the time the invention was made, it would have been obvious to one of ordinary skill in the art to implement resolution to resolve ambiguity of Barton. One of ordinary skill in the art would have been motivated to do this since it enables back and forth communication to resolve ambiguity. (Page 5 [0048-0049], Page 6 [0059] and Page 7 [0067-0068])

The Patent Office has made an assertion that Barton teaches "the receiver is for receiving a request message over the wireless link that requests additional features and the processor is automatically responsive to the request message to extract a second set of features from the digital version of the media sample and the transmitter is further to transmit the extracted second set" and that the teaching for this is found in paragraphs 0048, 0049, 0059, 0067, and 0068 of Barton. These five paragraphs are reproduced below:

[0048] Referring again to FIG. 1, the experiential environment sample is received by recognition engine 110 on line 117. Recognition engine 110 derives characteristics of the received sample by using data stored in database 115. Recognition 110 and database 115 are operationally coupled via line 119, as shown in FIG. 1. A variety of derivation methods may be used. In the case of audio samples, the techniques described in Appendix may be used. However, it is noted that the derivation methods that may be used in this invention are not limited to such techniques. The particular derivation method chosen is only required to be able to derive sufficient characteristics from **the experiential environment sample** to enable a predetermined event to be triggered. Thus, the strength or resolution of the derivation is coupled, in large measure, to the level of discrimination required in selecting an event to be triggered. As the number of potentially triggered events increases, the necessity to resolve ambiguity in the sample also increases.

[0049] For example, in the case of the exemplary embodiment where song lyrics corresponding to a broadcast song are sought by a user, a relatively large number of characteristics about **the sample** may be derived and compared against stored data to be able to identify the particular song from the many such songs that may be stored. That is, as more songs are potentially identified, more lyric delivery events are potentially triggered. By comparison, in se vice offerings where are relatively small number of events are potentially triggered, fewer sample characteristics need typically be derived in order to resolve ambiguity as to which event to trigger. Such service offering may include those where a binary "Yes" or "No" event may be triggered as may be the case for customer surveys and voting/polling type services.

[0059] The friends are prompted in the message or call to try to "Name that Tune" by identifying the song's title or artist from the small excerpt. The friend's guesses may be collected by the service provider using a variety of methods, including for example, an interactive web-site, telephone call center, email, or conventional mail. If no one correctly identifies the song, the song excerpt may be increased in length, or a different excerpt may be furnished, in an iterative manner, until a "winner" is determined.

[0067] Block 174 in FIG. 1 shows that control events may also be triggered in response to a sampled experiential environment in accordance with the invention. Control events are those that provide the user with an ability to control or otherwise manipulate information and data, services, or other events in a predetermined manner according to the captured sample received by a service provider. For example, a human resources recruiter may organize a data archive of job candidates and associated demographic data by engaging a service provider that automatically manipulates the data according to web-site images of potential hiring companies that are captured in a frame grabber running on the user's computer and uploaded to the service provider. In such cases, the candidate database can be sorted according the to captured web-document and derived by deriving preselected characteristics such as industry type, key-words in the text elements of the page, and other characteristics.

[0068] Communication events may be triggered in accordance with the invention as depicted by block 175 in FIG. 1. Communication events include, for example, communicative interactions among users, between users and the service provider, or such interactions between users, the service provider, and third parties.

Paragraph 0059 of Barton discloses "the song excerpt may be increased or a different excerpt may be furnished" for a group of friends in a game after the recognition engine 110 identifies the song. Preceding paragraph 0058 of Barton discloses "A game type entertainment event is then triggered by the service to automatically send a small excerpt of the originally recorded song (i.e., not the captured sample of the song) to a pre-determined group of the user's friends via" which clearly shows that Barton does not send a message to the capture device to send a second sample but, rather, selects another portion or excerpt from the captured sample of the song. That is, the song naming game of paragraph 0059 is an illustration of an application when a song has been identified through Barton's invention. Barton's recognition engine 110 does not recursively inquire of the capture device 102 for more information. It seems that there is but one sample passed from the capture device 102 to the recognition engine 110, which sample is then identified from the database 115 associated with the recognition engine. Also, it is noteworthy that whereas Barton performs any feature extraction in a remote device (i.e., the recognition 110) in contrast Wang discloses all feature extraction occurs on the client end (column 8, lines 13-21); polar opposite approaches.

Barton, like Wang, does not teach an apparatus configured to extract lower level but not higher level features that are transmitted off-apparatus for identification of a media from a media sample and does not teach a remote service configured to identify from

14

received lower level features and configured to extract higher level features from the received lower level features to uniquely identify a media corresponding to the media sample.

Vetro is apparently cited by the Patent Office for a teaching of high-level description schemes (col. 4, lines 44-46) in addition to a low-level representation (col. 4, lines 34-38) and SummaryDS (col. 22, lines 30-33). Vetro relates to "delivery systems that adapt information to available bit rates of a network" (col. 1, lines 15-17).

The relevance of Vetro is not understood in light of the currently pending claims. Any higher level features derived by Vetro are passed on as content information CI 302 to the CND manager 330 (column 8, line 60, through column 9, line 9) which is used to determine an optimal transcoding strategy for switchable transcoder 340.

Vetro, like Wang and Barton, does not teach an apparatus configured to extract lower level but not higher level features that are transmitted off-apparatus to identify a media from a media sample and does not teach a remote service configured to identify from received lower level but not higher level features and configured to extract any needed higher level features from the received lower level but not higher level features to uniquely identify a media corresponding to the media sample.

Thus, claims 1-9, 12, 16-20, 22-27, 30, 35, 37, 38, 40, and 47-50 are allowable over these three references, alone or in combination.

New claims 51 to 57 are believed to be allowable over the cited three references, alone or in combination.

The Patent Office is respectfully requested to reconsider and remove the rejections of the claims under 35 U.S.C. 103(a) based on Wang, Barton, and Vetro, alone or in combination, and to allow all of the pending claims 1-9, 12, 16-20, 22-27, 30, 35, 37, 38, 40, and 47-56 as now presented for examination. An early notification of the allowability of claims 1-9, 12, 16-20, 22-27, 30, 35, 37, 38, 40, and 47-56 is earnestly solicited.

Respectfully submitted:

Walter J. Malinowski
Reg. No.: 43,423

October 7, 2008
Date

Customer No.: 29683
HARRINGTON & SMITH, PC
4 Research Drive
Shelton, CT 06484-6212

Phone: (203) 925-9400, extension 19
Email: wmalinowski@hspatent.com
Facsimile: (203) 944-0245

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

10/7/2008
Date

Name of Person Making Deposit

16